Built-in Self Repair for SRAM Array using Redundancy

A. Padma Sravani, Dr. M. Satyam
Centre for VLSI and Embedded System Technologies
International Institute of Information Technology, Hyderabad, India
Email: padmasravani.a@research.iiit.ac.in, satyam@iiit.ac.in

***Abstract*—In this paper, a built-in self repair technique for word-oriented two-port SRAM memories is presented. The technique is implemented by additional hardware design instead of traditional software diagnostic procedures and the computation time is minimized. A built-in self-test (BIST) is used to detect the faulty locations which are isolated immediately after detection. Therefore, the redirection process can be executed as soon as possible. Spare rows are used to replace the faulty rows. The hardware overhead of the automatic fault isolation design depends on size of memory system. All the repairs using BISR circuit are done at power on.**

***Index Terms*—SRAM, two-port memories, PVT faults, fault isolation, BISR**

I. Introduction

Today's data-dominated multimedia applications require more memory than ever before. On-chip SRAM memories begin to dominate the chip area and have become the focus of technology scaling. However, the physical limitations of the technology scaling jeopardize further progress of microelectronics as scaling results in process variations. Increase in process variations causes parametric variations in transistor feature sizes and threshold voltages due to random dopant fluctuations, line edge roughness, sub-wavelength lithography[5] [6]. Closely matched devices and small transistor sizes which matter the most when designing SRAM memories, are the first to suffer from the side-effects of scaling. Random nature of local process variation causes defects to have random and uniform distribution [8] [9] .This adversely affects the expected system yield. Since memory is one of the biggest blocks of any system, it is more prone to faults under process variation. A failure in memory cell can occur due to *i)* an increase in cell access time and *ii)* unstable read/write operation *iii)* inability to hold the cell content [11]. The mismatch in device parameters will increase the probability of these failures.

As integrated circuits are growing rapidly in component density and scale, error detection and fault isolation have become more difficult and expensive. The decrease in cost of manufacturing as against increase in cost of testing has led to various approaches to design for diagnosability [l] [2] and design for testability [3] [4]. Methods to improve reliability by providing internal fault diagnosis/tolerance with additional hardware are gaining importance. On the other hand, the introduced test area overhead must be kept as low as possible. In particular, a duplication of the entire memory is not allowed. Many techniques have been proposed to handle failing cells in SRAM structures. Fault masking methods like Triple Modular Redundancy (TMR) [7] can be used for small size memories. .However, such techniques are not feasible for high capacity memories because of large additional hardware overhead.

Statistical sizing and optimization of the SRAM cell for yield enhancement is suggested in [10]. This pre-silicon technique could improve production yield, however, it is limited by conflict in sizing requirement for different types of failures [15]. Conventional procedures have faulty lines replaced with redundant lines by switching between decoders [13] [12]. Access time penalty is inevitable in this method, which is not desirable for high-speed SRAMs.

More recent approaches use BISR concepts. In [14], faulty address and its data are stored in the redundancy logic requiring address comparison in BISR, causing extra power during the normal operation. In [15], BISR circuit requires fail address memory (consists of CAM) and a spare memory. This method also compares input addresses with stored addresses during normal operation increasing power consumption. Dedicated CAM structure is largest part of hardware overhead.

We propose Built-in Self Repair SRAM architecture which uses BIST (Built-in Self Test) to detect and replace faulty rows using spare rows. BISR circuit which operates only during the power on is proposed. During normal operation, address comparison with faulty addresses is not required, because the necessary repair is done during power on. This architecture has minimum additional hardware overhead and access time penalty. The repair capability of SRAM depends on the number of spare rows used. This fault isolation and redirection to spare rows scheme can be completely implemented with additional hardware. Here we assume that the additional hardware is fault-free because of its relative simplicity.

The paper is organized as follows. In section II, review of 8t - SRAM cells and two-port memories is given. Section III describes different failures in SRAM cell. Section IV explains the BIST architecture. Section V shows the redirection of faulty rows to spare rows. Section VI gives the simulation results and finally in section VII we draw conclusions.

DOI: 01.IJCom.02.01.123

II.Review of SRAM cells

The conventional static random access memory (SRAM) cell consists of 6 transistors (6T). To meet the expected demands of parallel or pipelined microprocessors and for increased data throughput multi-port SRAMs are often used. Multi-port SRAMs are implemented using SRAM due to its fast operation and the ability to support multiple read and write operations simultaneously. A multi-port memory, such as two-port memory as shown in fig. 1 is actually a single memory cell with two entirely independent sets of data, address and control lines. Such a memory can be written or read from two different paths. By using two-port SRAMs, the efficiency of the memory accesses can be doubled. In the two-port memories, the read and write ports are completely independent and can access any location simultaneously (but simultaneous access to same cell by both ports cannot happen). In this paper, we study the faults occurring in the two-port SRAM cells. The memory is word-oriented and is modelled at the transistor level. The circuit operation is studied during presence of fault. A set of test patterns are applied to detect the faults. A failure in any of the cells in a row of the memory will make that row faulty and entire row is isolated and then redirected.

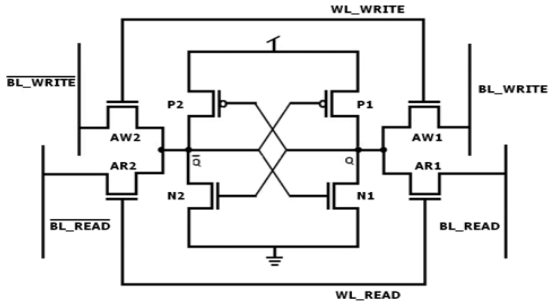

Fig 1. Two-port SRAM cell

III.FAILURE MECHANISM IN SRAM

Process variations in SRAM cells cause one of the following failures - *access failure, read failure, write failure and hold-stability failure.* Read failure and hold-stability failure occur in the presence of excessive variations in device parameters coupled with increased disturbance to the cell and very low supply voltages. The dynamic stability of the cell during the read and write operation is defined by the robustness of the cell against threshold voltage variations. Write failure and access failure may result even in the presence of slight variations. Employing aggressive timing and low supply voltages leads to these failures which have more concern.

IV.BIST ARCHITECTURE

Built In Self Test allows performing self test by designing and integrating additional hardware and software features in memory arrays. BIST makes the testing of a chip easier, faster, more efficient, and less costly. The Built-in self test is done at every power on. In the proposed architecture as shown in fig. 2 the $T/\overline{N}$ signal initiates the testing process. The multiplexers isolate all the SRAM inputs and select either normal system inputs or test inputs generated by the test pattern generator. Theselection choice is done with $T/\overline{N}$ input which is used to select test input address and test input data. In this design it is assumed a BIST controller is present which generates the $T/\overline{N}$ signal and test pattern generator which generates necessary inputs for SRAM array. The BIST controller will signal the end of testing process by making the signal $T/\overline{N}$ low.

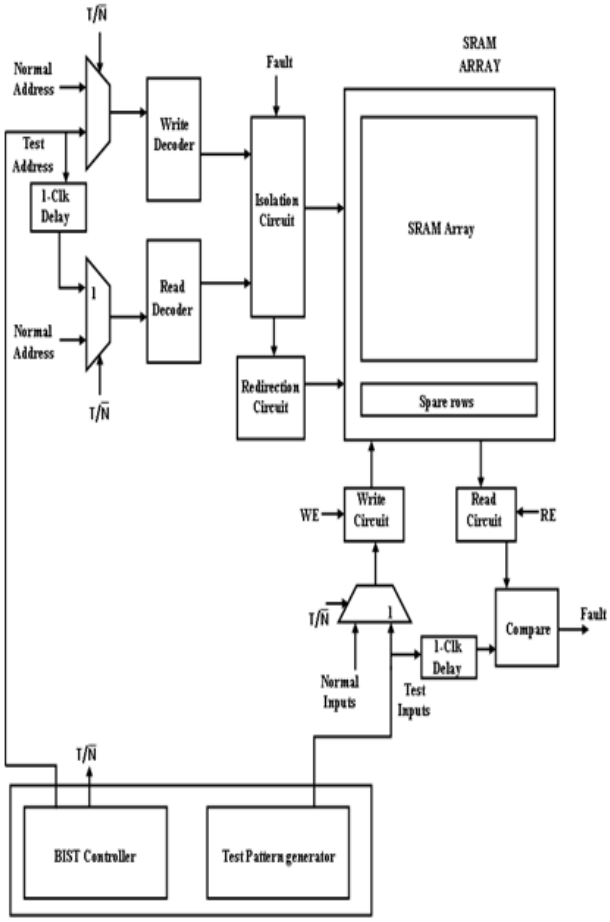

Fig 2. SRAM array with BIST and comparison circuit, isolation circuit and redirection circuit

The test results are collected in the isolation circuit block of the design, which determine whether the selected row in SRAM array is fault-free or faulty. It also generates necessary signals to redirect those lines to the spare array rows. Fig. 2 shows the entire architecture used in this design.

V.FAULT DETECTION AND REDIRECTION CIRCUITS

a) Fault Detection

A comparison circuit is used to detect the faults in SRAM. The fault in SRAM can occur due to any one of the above reasons discussed in Section III. But the fault is detected only when a row is read. In this design a row of SRAM array is written and is read in the following cycle.

The input data and the data read during read operation are compared to detect the fault. A bit-wise

DOI: 01.IJCom.02.01.123

comparison between input data (which should be delayed for 1-Clk) and sense amplifier outputs is done. An error in single SRAM cell will make the entire row faulty. The generated *fault* signal determines the fault status of the row.

b) Fault isolation of the failing row

When *fault* = 1, the row is detected as faulty. This *fault* signal is directed to all the isolations circuits. To mark the particular faulty row it is ANDed with WL_read (the row from which the read operation occurred). The resulting signal, from the isolation circuit module will mark the row as faulty through the *fault* signal. The pointers are updated to invalidate the row. The isolation of row would mean the wordlines for write and read of that row are permanently grounded and the row can never be accessed. The disconnect logic is implemented using switches. The hardware implementation of isolation circuit is shown in fig. 3.

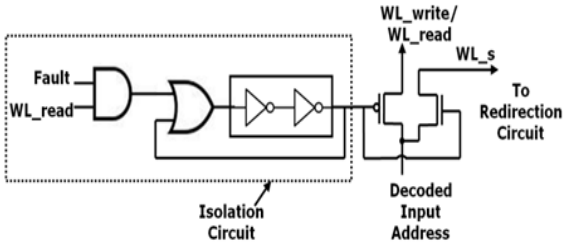

Fig. 3. Isolation circuit used to isolate the faulty rows

For each wordline the above shown isolation circuit will be used and the output of it will control the switches which direct the decoded input address to normal row (in fault-free condition) or spare row (in faulty condition). The fault status of the row is stored in the inverters and can be used for further analysis.

c) Repair using Redundancy

Redundancy is to provide multiple identical instances of the same system and switching to one of the remaining instances in case of a failure. When a failure occurs, the system must be able to isolate the failed component. This requires the addition of dedicated failure detection mechanisms and redirection logic to redirect to spare locations.

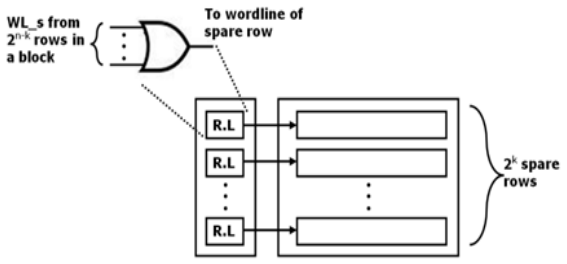

Fig. 4 Redirection circuit to redirect faulty locations to spare locations

Consider an SRAM array with 2^n rows and 2^k spare rows where k << n. The 2^n rows are divided into 2^k blocks. Any fault in one of the rows in those 2^k blocks (each of size 2^{n-k}) will be redirected to one of the 2^k spare locations. It is assumed the faults are normally distributed and during each consideration, only one faulty row occurs within a block and it can be redirected to one of the 2^k spare locations.

Fig. 4 shows the circuit implementation of redirecting faulty locations to spare locations. The redirection logic (R.L) will redirect the faulty row address to one of the spare locations. The design has minimum hardware overhead and the access delay due to additional hardware is also kept low.

VI. Simulation results

All the simulations were carried out in HSPICE with 1V supply voltage using the Predictive Technology Model (PTM) [16] files at 65nm technology node.

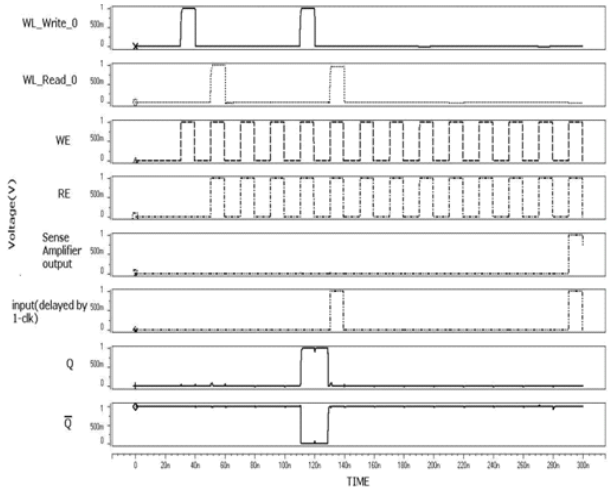

Fig. 5 Read failure of a cell causing incorrect read operation (incorrect sense amplifier output)

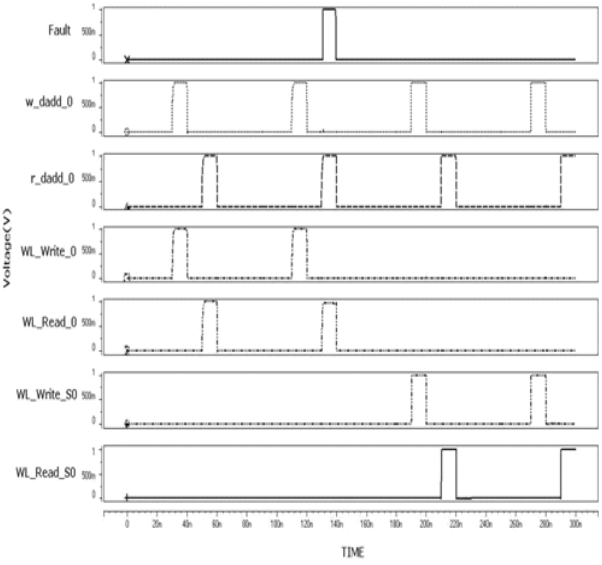

Fig. 6 The fault signal and wordlines before and after detection of fault

For the simulation a fault is introduced in bit#0 of row#0 in SRAM array. w_dadd_0 and r_dadd_0 are the address decoder outputs for write port and read port of row#0. When they are high, the respective wordlines, WL_Write_0 and WL_Read_0 are made high and the row is accessed for write and read respectively. During second time read of

DOI: 01.IJCom.02.01.123

row#0 the data in the cell (Q and $\overline{Q}$) is flipped causing read failure. The sense amplifier output is wrong and does not match with the delayed input. Hence row is found faulty and the *fault* signal goes high, indicating the fault. It can be observed that in the following cycles for the valid input address for row#0, the WL_Write_0 and WL_Read_0 remain low i.e., the row contents are not accessed, but instead the wordlines (for write and read ports) of spare row are raised high and it is used.

VII.CONCLUSION

A built-in self repair for word-oriented SRAM is proposed. The fault detection is done using built-in self test and comparison logic. The isolation circuit has isolated the faulty rows from the array. Repair is achieved by redirecting faulty rows to the spare rows using redirection circuit.

REFERENCES

[1] H. Ma and Y. Liu, "Design for diagnosable multiple-output digital systems," in Proc. IEEE VLSI Test Symposium, IEEE Computer Society Press, Atlantic City, NJ, April 15 -17,1991, pp.204-209.

[2] H. Ma and Y. Liu, "Efficient placement of error checkers in VLSI systems," in Proc. Test Engineering Conference, Atlanta, GA, June 25-27, 1991, pp.103-111.

[3] H. Ma and T. N. Raiashekhara, "A testable design of semiconductor RAMS," d Roc. The 30th Midwest Conference on Wts and Systems, Syracuse., NY, 1987, pp.540-543.

[4] T. W. Williams and K P. Parker, "Design for testability – a survey," IEEE Trans. Comput., C-31, No.1, Jan. 1982, pp.2-15.

[5] A. Bhavanagarwala, X. Tang, and J. Meindl, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE J. Solid State Circuits*, vol. 36, no. 4, pp. 658-665, Apr. 2001.

[6] G. Gielen et al., "Analog and digital circuit design in 65 nm CMOS: end of the road?", in *Proceedings of the Design, Automation and Test in Europe*, 2005

[7] D.P. Siewiorek and R.S. Swarz, "The Theory and Practice of Reliable System Design", Digital Press, 1982

[8] S. R. Nassif "Modeling and Analysis of manufacturing variation" in Proc. CICC, 2001 pp/ 223-228

[9] S. Borkar, T. Karnik, et al., "Process Variation and impact on circuits and micro architectures," in Proc DAC 2003 pp338-342

[10] S. Mukhopadhyay, et al., "Statistical design and optimization of SRAM cell for yield enhancement," in Proc. Int. Conf. Computer Aided Design (ICACD), Nov. 2004, pp. 10–13.

[11] S. Mukhopadhyay, H. Mahmoodi, and K. Roy, "Modeling and estimation of failure probability due to parameter variations in nano-scale SRAM's for yield enhancement," in Proc. VLSI Circuit Symp., Jun. 2004, pp. 64–67

[12] Stanley Schuster,"Multiple Word/Bit Line Redundancy for Semiconductor Memories", IEEE Journal of Solid State Circuits, Vol.SC-13, No.5, October 1978, pp.698-703.

[13] T. Mano and M. Wada and N. Ieda, M. Tanimoto, "A Redundancy Circuit for a Fault-Tolerant 256K MOS RAM", IEEE Journal of Solid State Circuits, Vol.SC-17, No.4, August 1982, pp.726-731.

[14] V. Schober and S. Paul and O. Picot, "Memory Built-In Self-Repair using redundant word", IEEE International Test Conference, 2001, pp.995-1001.

[15] A. Tanabe and T. Takeshima and H. Koike and Y. Amimoto and M. Takada and et. al, "A 30-ns 64-Mb DRAM with Built-in Self-Test and Self- Repair Function", IEEE Journal of Solid State Circuits, Vol.27, No.11, November 1992, pp.1525-1533.

[16] Berkeley Predictive Technology Model website, http://www.eas.asu.edu/~ptm/

DOI: 01.IJCom.02.01.123